Fusion of pixel-based and object-based features for classification of urban hyperspectral remote sensing data

Wenzhi liao[a,*], Frieke Van Coillie[b], Flore Devriendt[b], Sidharta Gautama[a], Aleksandra Pizurica[a], Wilfried Philips[a]
[a] iMinds-Telin-IPI, Faculty of Engineering and Architecture,
[b] FORSIT Research Unit, Faculty of Bioscience Engineering,
Ghent University, Belgium
*Corresponding author: wenzhi.liao@telin.ugent.be, +3292643270, +3292644295

Abstract: Hyperspectral imagery contains a wealth of spectral and spatial information that can improve target detection and recognition performance. Typically, spectral information is inferred pixel-based, while spatial information related to texture, context and geometry are deduced on a per-object basis. Existing feature extraction methods cannot fully utilize both the spectral and spatial information. Data fusion by simply stacking different feature sources together does not take into account the differences between feature sources. In this paper, we propose a feature fusion method to couple dimension reduction and data fusion of the pixel- and object-based features of hyperspectral imagery. The proposed method takes into account the properties of different feature sources, and makes full advantage of both the pixel- and object-based features through the fusion graph. Experimental results on classification of urban hyperspectral remote sensing image are very encouraging.
Keywords: Data Fusion, Remote Sensing, Object-based features, Pixel-based features.

1. Introduction

Nowadays, advanced sensor technology and image processing algorithms allow us to measure different aspects of the objects on the Earth's surface. Both spectral and spatial characteristics can be extracted from hyperspectral remote sensing data, and used to identify and discriminate between ground objects. In particular, spectral features inferred from hyperspectral image pixels provide a detailed description of the spectral signatures of ground covers, whereas spatial features deduced from image objects give detailed information about texture, context and geometry of the surface objects (Rastner et al., 2014; Myint et al., 2011; Weih and Riggan, 2010).

It is clear that spectral or spatial features alone might not be sufficient to obtain reliable classification results in an urban context. Instead, the combination of both data sources can contribute to a more comprehensive interpretation of the ground objects (Zhang and Tang, 2013; Chan et al., 2009). For example, spectral signatures cannot differentiate between objects made of the same material (e.g. roofs and roads made with the same asphalt), while the latter may be easily distinguished by their geometry. On the other hand, spatial features alone may fail to discriminate between objects that are quite different in nature (e.g. grass field and swimming pool), but similar in their geometry.

Stacking multi-source features together is a widely applied data fusion technique for classification (Chan et al., 2009). These methods first apply feature extraction on each individual feature source, after which all features are concatenated into one stacked vector for classification. While such methods are appealing due to their simplicity, they do not always perform better (and sometimes worse) than using a single feature source.

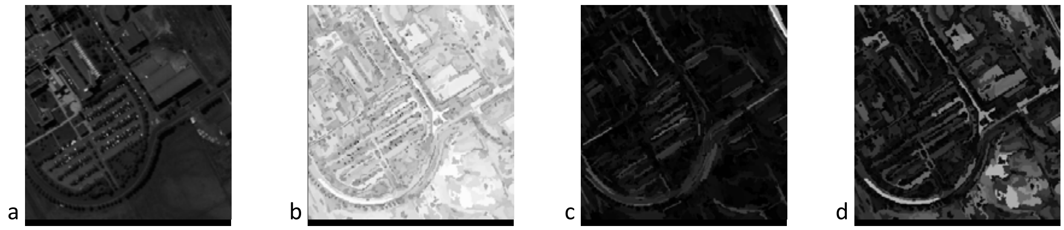


Figure 1. Some of the object-based features, a. First PC extracted from the original HS image, b. GLCM entropy (all directions) (PC1), c. Length-to-width ratio, and d. Area based on sub-object analysis.

This is because the value of different components in the stacked feature vector can be significantly unbalanced. As a consequence, the information contained by different feature sources is not equally represented or measured. Furthermore, by stacking several feature sources together, the pooled data may contain redundant information. Last but not least, the increase in the dimensionality of the stacked features, combined with the limited number of labelled samples, may together lead to the problem of the "curse of dimensionality".

Therefore, we propose a graph-based fusion method to couple dimension reduction and data fusion of the pixel- and object-based features. First, object-based features are generated on the first few principal components of the original hyperspectral data. Second, we build a fusion graph where only the feature points with similar spectral and spatial characteristics are connected. Finally, we solve the problem of multi-source feature fusion by projecting both feature sources into a linear subspace, on which neighborhood data points (i.e. with both similar spectral and spatial characteristics) in the high-dimensional feature space are kept on neighborhood in the low-dimensional projected subspace as well. This way, the proposed method takes into account the properties of different feature sources and makes full advantage of both pixel- and object-based features through the fusion graph. Our graph-based data fusion method won the ``Best Paper Challenge'' award of 2013 IEEE Data Fusion Contest, but with experiments on multi-sensor data sources (hyperspectral and LiDAR data) (Debes et al., 2014).

2. Hyperspectral data and object-based features

Experiments were run on hyperspectral remote sensing data from urban area. The data were collected by the ROSIS sensor over the University of Pavia, Italy, with 610×340 pixels and 103 spectral bands after removal of noisy bands. This data set includes 9 land cover/use classes, with very fine spatial resolution of 1.3 meters by pixel.

To obtain object-based features from a hyperspectral image, principal component analysis (PCA) was first applied to the original hyperspectral data, and the first 4 PCs were selected (representing 99% of the cumulative variance) to generate the object-based features. A multi-resolution segmentation was performed with eCognition® on the first 4 PCs of the original HS image, and a total of 28 object-based features were generated: Mean, StdDev, Ratio, GLCM Contrast, Entropy, Correlation in the first 4 PCs, respectively; Area, Length/width, Shape index and Width. Fig.1 shows some of the object-based features.

3. Proposed feature fusion method

We first build a graph for each feature source, for example, the graph constructed by spectral features (i.e., $G^{Spe}=(X^{Spe}, A^{Spe}))$, where X^{Spe} denotes pixels from original HS data, and A represents the edges of the graph. The edge between data point x_i^{Spec} and x_j^{Spec} is here

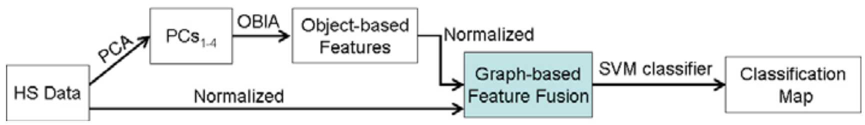


Figure 2. Proposed data fusion framework.

denoted as $A_{ij}^{Spec} \in \{0, 1\}$; $A_{ij}^{Spec}=1$ if x_i^{Spec} and x_j^{Spec} are "close" and $A_{ij}^{Spec}=0$ if x_i^{Spec} and x_j^{Spec} are "far apart". The "close" is defined as belonging to *k* nearest neighbors (*k*NN) of the other data points. The *k*NNs of the data point x_i^{Spec} are its *k* nearest neighbors in terms of spectral signatures. On the other hand, when the graph is constructed by object-based features, the *k*NNs of the data point x_i^{Spat} are its *k* nearest neighbors in terms of spatial characteristics. We define a fusion graph $G^{Fus}=(X^{Sta}, A^{Fus})$, where $X^{Sta}=[X^{Spec}; X^{Spat}]$, $A^{Fus} = A^{Spec} \odot A^{Spat}$. The operator $\odot$ denotes element-wise multiplication, i.e. $A_{ij}^{Fus}= A_{ij}^{Spec} A_{ij}^{Spat}$. This means that the stacked data point x_i^{Sta} and x_i^{Sta} are connected only if they have similar both spectral and spatial characteristics. Fig. 2 shows our proposed fusion framework, for more details to obtain the fused features, we refer the readers to our recent work (Debes et al., 2014).

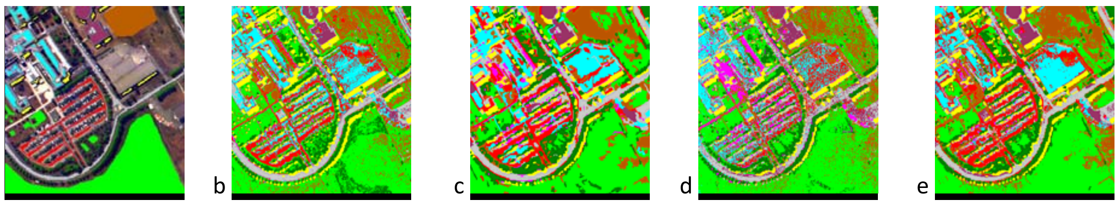
4. Experimental results

The SVM classifier with radial basis function (RBF) kernels is applied in our experiments. We compare our proposed method with the schemes of (1) using original HS data (Spectral); (2) using the object-based features computed on the first 4 PCs of the original HS data (OBIA); (3) stacking all spectral and object-based features together (Sta); (4) features fused by using the graph constructed by stacked features (i.e. LPP (He and Niyogi, 2004)) (LPP). We randomly select 20 labelled samples per class for training. The classifiers were evaluated against the remaining labelled samples by measuring the Overall Accuracy (OA), the Average Accuracy (AA) and the Kappa coefficient (κ). Tab. 1 shows the accuracies obtained from the experiments, and Fig. 3 shows the classification maps.

Our proposed feature fusion method yields better overall performance, with 4% improvement compared to using single pixel-/object-based features, and with almost 3% improvements over the other fusion schemes in terms of OA. The use of pixel-based features alone yields OA of less than 75%, and thus produces a noisy classification map. The OBIA approach can produce much better results, with more 10% improvements in OA and much smoother classification map. The Sta method produced lower accuracies than only using single object-based features, indicating that the spatial information contained in the original HS data was not well exploited in such a stacked architecture. This is not surprising because values of different feature components can be significantly unbalanced, and thus the information contained by different features is not equally well represented. The same problems happen when using the stacked features to build a graph in the LPP method. The proposed method overcomes these problems and achieves thereby better classification accuracies.

Table 1. Classification Accuracy assessment report.

	Spectral	OBIA	Sta	LPP	Proposed
No. of Features	103	28	131	26	24
OA (%)	74.91	86.19	81.39	87.05	90.66
AA (%)	83.13	87.99	84.20	89.12	91.76
κ	0.685	0.820	0.759	0.837	0.878



a b c d e

Figure 3. Part of classification maps, a. Ground truth, b. Spectral, c. OBIA, d. Sta, and e. Proposed method.

5. Conclusions

In this paper, we presented a graph-based feature fusion method to include both pixel- and object-based features in the classification process. Experiments on a hyperspectral image demonstrate that data fusion by simply stacking several feature sources together may perform worse than using a single feature source. The proposed method considers the differences between pixel- and object-based features, makes full advantage of both feature sources through fusion graph, and improves thereby strongly the classification performance.

Acknowledgment

This work was financially supported by the SBO-IWT project Chameleon: Domain-specific Hyperspectral Imaging Systems for Relevant Industrial Applications. The authors would like to thank Prof. Paolo Gamba from the University of Pavia, Italy, for kindly providing the University Area hyperepsectral data.

References

Rastner P., Bolch T., Notarnicola C., Paul F., 2014, A Comparison of Pixel- and Object-Based Glacier Classification With Optical Satellite Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Volume 7, Issue 3: 853-862.

Myint S. W., Gober P., Brazel A., Clarke S. G., Weng Q., 2011, Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment, Volume 115, Issue 5:1145-1161.

Zhang A., Tang P., 2013, Fusion algorithm of pixel-based and object-based classifier for remote sensing image classification. 2013 IEEE International Geoscience and Remote Sensing Symposium, Melbourne, Australie, pp. 2740-2743.

Chan J.C.W., Bellens R., Canters F., Gautama S., 2009, An assessment of geometric activity features for per-pixel classification of urban man-made objects using very high resolution satellite imagery. Photogrammetric Engineering & Remote Sensing, Volume 75, Issue 4: 397-411.

Debes C., Merentitis A., Heremans R., Hahn J., Frangiadakis N., Kasteren T., Liao W., Bellens R., Pizurica A., Gautama S., Philips W., Prasad S., Du Q., Pacifici F., 2014, Hyperspectral and LiDAR Data Fusion: Outcome of the 2013 GRSS Data Fusion

Contest. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (In Press).

Weih R. C., Riggan N. D., 2010, Object-Based Classification vs. Pixel-Based Classification: Comparative Importance of Multi-Resolution Imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (GEOBIA 2010), Volume: XXXVIII-4/C7.

He X. F., Niyogi P., 2004, Locality preserving projections. Advances in Neural Information Processing Systems 16, MIT Press, Cambridge, pp. 153-160.